What is claimed is:

1.   A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a first frame removing unit removing redundant video frames from the given video frames;

a second frame removing unit removing video frames that do not contain a text area from the given video frames;

a third frame removing unit detecting and removing redundant video frames caused by image shifting from the given video frames; and

an output unit outputting remaining video frames as candidate text change frames.

2.   The text change frame detection apparatus according to claim 1, wherein the first frame removing unit includes:

an image block validation unit determining whether two image blocks in the same position in two video frames of the given video frames are a valid block pair that has an ability to show a change of image contents;

an image block similarity measurement unit calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar; and

5   a frame similarity judgment unit determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs, and the first frame removing unit removes a similar

10   video frame as a redundant video frame.

3.   The text change frame detection apparatus according to claim 1, wherein the second frame removing unit includes:

15   a fast and simple image binarization unit generating a first binary image of a video frame of the given video frames;

a text line region determination unit determining a position of a text line region by

20   using a horizontal projection and a vertical projection of the first binary image;

a rebinarization unit generating a second binary image of every text line region;

a text line confirmation unit determining

25   validity of a text line region by using a

difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region; and

5　　　　a text frame verification unit confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

10

4.　The text change frame detection apparatus according to claim 1, wherein the third frame removing unit includes:

a fast and simple image binarization unit

15 generating binary images of two video frames of the given video frames;

a text line vertical position determination unit determining a vertical position of every text line region by using horizontal projections of the

20 binary images of the two video frames;

a vertical shifting detection unit determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation

25 between the horizontal projections; and

a horizontal shifting detection unit determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation

5   between vertical projections of every text line in the binary images of the two video frames, and the third frame removing unit removes a similar video frame as a redundant video frame caused by the image shifting.

10

5.   A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

15   an image block validation unit determining whether two image blocks in the same position in two video frames of given video frames are a valid block pair that has an ability to show a change of image contents;

20   an image block similarity measurement unit calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar;

a frame similarity judgment unit determining

25   whether the two video frames are similar by using a

ratio of a number of similar image blocks to a total number of valid block pairs; and

an output unit outputting remaining video frames after a similar video frame is removed, as

5   candidate text change frames.

6.   A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus

10   comprising:

a fast and simple image binarization unit generating a first binary image of a video frame of the given video frames;

a text line region determination unit

15   determining a position of a text line region by using a horizontal projection and a vertical projection of the first binary image;

a rebinarization unit generating a second binary image of every text line region;

20   a text line confirmation unit determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a

25   total number of pixels in the text line region;

a text frame verification unit confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the

5   set of continuous video frames; and

an output unit outputting remaining video frames after the non-text frames are removed, as candidate text change frames.

10   7.   A text change frame detection apparatus that selects a plurality of video frames including text contents from given video frames, said apparatus comprising:

a fast and simple image binarization unit

15   generating binary images of two video frames of the given video frames;

a text line vertical position determination unit determining a vertical position of every text line region by using horizontal projections of the

20   binary images of the two video frames;

a vertical shifting detection unit determining a vertical offset of image shifting between the two video frames and a similarity of the two video frames in a vertical direction by using correlation

25   between the horizontal projections;

a horizontal shifting detection unit determining a horizontal offset of the image shifting and a similarity of the two video frames in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and

an output unit outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

8. A text extraction apparatus that extracts at least one text line region from a given image, said apparatus comprising:

an edge image generation unit generating edge information of the given image;

a stroke image generation unit generating a binary image of candidate character strokes in the given image by using the edge information;

a stroke filtering unit removing a false stroke from the binary image by using the edge information;

a text line region formation unit combining a plurality of strokes into a text line region;

a text line verification unit removing a false character stroke from the text line region and

reforming the text line region;

a text line binarization unit binarizing the text line region by using a height of the text line region; and

5   an output unit outputting a binary image of the text line region.

9.   The text extraction apparatus according to claim 8, wherein the edge image generation unit 10   includes:

an edge strength calculation unit calculating edge strength for every pixel in the given image by using a Sobel edge detector;

a first edge image generation unit generating 15   a first edge image by comparing the edge strength of every pixel with a predefined edge threshold and setting a value of a corresponding pixel in the first edge image to one binary value if the edge strength is greater than the threshold and the 20   other binary value if the edge strength is less than the threshold; and

a second edge image generation unit generating a second edge image by comparing the edge strength of every pixel in a window centered at a position 25   of every pixel of the one binary value in the first

edge image with mean edge strength of the pixels in the window and setting a value of a corresponding pixel in the second edge image to the one binary value if the edge strength of the pixel is greater

5   than the mean edge strength and the other binary value if the edge strength of the pixel is less than the mean edge strength.

10.   The text extraction apparatus according to
10   claim 9, wherein the stroke image generation unit includes a local image binarization unit binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the candidate character strokes by using a
15   window centered at a position of every pixel of the one binary value in the second edge image.

11.   The text extraction apparatus according to claim 9, wherein the stroke filtering unit
20   includes:
        a stroke edge coverage validation unit checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels of the one binary value in the
25   second edge image, determining that the stroke is a

valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold, and removing the invalid stroke; and

5   a long straight line detection unit removing a large stroke by using a width and a height of the stroke.

12. The text extraction apparatus according to
10 claim 9, wherein the text line binarization unit includes:

an automatic size calculation unit determining a size of a window for binarization; and

a block image binarization unit binarizing a
15 gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the text line region by using the window centered at a position of every pixel of the one binary value in the second edge image.

20

13. The text extraction apparatus according to claim 8, wherein the text line region formation unit includes a stroke connection checking unit checking whether two adjacent strokes are
25 connectable by using an overlap ratio of heights of

the two strokes and a distance between the two strokes, and the text line region formation unit combines the plurality of strokes into a text line region by using a result of checking.

14.	The text extraction apparatus according to claim 8, wherein the text line verification unit includes:

	a vertical false stroke detection unit checking every stroke with a height higher than a mean height of strokes in the text line region, and marking the stroke as a false stroke if the stroke connects two horizontal text line regions into one big text line region;

	a horizontal false stroke detection unit checking every stroke with a width larger than a threshold determined by a mean width of the strokes in the text line region, and marking the stroke as a false stroke if a number of strokes in a region that contains the stroke is less than a predefined threshold; and

	a text line reformation unit reconnecting strokes except for a false stroke in the text line region if the false stroke is detected in the text line region.

15. A text extraction apparatus that extracts at least one text line region from a given image, said apparatus comprising:

5   an edge image generation unit generating an edge image of the given image;

     a stroke image generation unit generating a binary image of candidate character strokes in the given image by using the edge image;

10   a stroke filtering unit checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels indicating an edge in the edge image, determining that the stroke is a valid stroke if the overlap

15 rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold, and removing the invalid stroke; and

     an output unit outputting information of

20 remaining strokes in the binary image of the candidate character strokes.

16. A computer-readable storage medium storing a program used to direct a computer, that selects a

25 plurality of video frames including text contents

from given video frames, to perform a process comprising:

removing redundant video frames from the given video frames;

5   removing video frames that do not contain a text area from the given video frames;

detecting and removing redundant video frames caused by image shifting from the given video frames; and

10   outputting remaining video frames as candidate text change frames.

17.   The storage medium according to claim 16, wherein the removing redundant video frames

15   includes:

determining whether two image blocks in the same position in two video frames of the given video frames are a valid block pair that has an ability to show a change of image contents;

20   calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar; and

determining whether the two video frames are similar by using a ratio of a number of similar

25   image blocks to a total number of valid block pairs,

and the removing redundant video frames removes a similar video frame as a redundant video frame.

18.   The storage medium according to claim 16,
5   wherein the removing video frames that do not contain the text area includes:

generating a first binary image of a video frame of the given video frames;

determining a position of a text line region
10   by using a horizontal projection and a vertical projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by
15   using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region; and

20   confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames.

25   19.   The storage medium according to claim 16,

wherein the detecting and removing redundant video frames caused by image shifting includes:

generating binary images of two video frames of the given video frames;

5 determining a vertical position of every text line region by using horizontal projections of the binary images of the two video frames;

determining a vertical offset of image shifting between the two video frames and a

10 similarity of the two video frames in a vertical direction by using correlation between the horizontal projections; and

determining a horizontal offset of the image shifting and a similarity of the two video frames

15 in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames,
and the detecting and removing redundant video frames removes a similar video frame as a redundant·

20 video frame caused by the image shifting.

20. A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents

25 from given video frames, to perform a process

comprising:

determining whether two image blocks in the same position in two video frames of given video frames are a valid block pair that has an ability to show a change of image contents;

calculating a similarity of two image blocks of the valid block pair and determining whether the two image blocks are similar;

determining whether the two video frames are similar by using a ratio of a number of similar image blocks to a total number of valid block pairs; and

outputting remaining video frames after a similar video frame is removed, as candidate text change frames.

21. A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

generating a first binary image of a video frame of the given video frames;

determining a position of a text line region by using a horizontal projection and a vertical

projection of the first binary image;

generating a second binary image of every text line region;

determining validity of a text line region by using a difference between the first binary image and the second binary image and a fill rate of a number of foreground pixels in the text line region to a total number of pixels in the text line region;

confirming whether a set of continuous video frames are non-text frames that do not contain a text area by using a number of valid text line regions in the set of continuous video frames; and

outputting remaining video frames after the non-text frames are removed, as candidate text change frames.

22. A computer-readable storage medium storing a program used to direct a computer, that selects a plurality of video frames including text contents from given video frames, to perform a process comprising:

generating binary images of two video frames of the given video frames;

determining a vertical position of every text

line region by using horizontal projections of the binary images of the two video frames;

determining a vertical offset of image shifting between the two video frames and a

5   similarity of the two video frames in a vertical direction by using correlation between the horizontal projections;

determining a horizontal offset of the image shifting and a similarity of the two video frames

10   in a horizontal direction by using correlation between vertical projections of every text line in the binary images of the two video frames; and

outputting remaining video frames after a similar video frame is removed, as candidate text

15   change frames.

23. A computer-readable storage medium storing a program used to direct a computer, that extracts at least one text line region from a given image, to

20   perform a process comprising:

generating edge information of the given image;

generating a binary image of candidate character strokes in the given image by using the

25   edge information;

removing a false stroke from the binary image by using the edge information;

combining a plurality of strokes into a text line region;

5   removing a false character stroke from the text line region and reforming the text line region;

binarizing the text line region by using a height of the text line region; and

10   outputting a binary image of the text line region.

24. The storage medium according to claim 23, wherein the generating edge information includes:

15   calculating edge strength for every pixel in the given image by using a Sobel edge detector;

generating a first edge image by comparing the edge strength of every pixel with a predefined edge threshold and setting a value of a corresponding

20   pixel in the first edge image to one binary value if the edge strength is greater than the threshold and the other binary value if the edge strength is less than the threshold; and

generating a second edge image by comparing

25   the edge strength of every pixel in a window

centered at a position of every pixel of the one binary value in the first edge image with mean edge strength of the pixels in the window and setting a value of a corresponding pixel in the second edge image to the one binary value if the edge strength of the pixel is greater than the mean edge strength and the other binary value if the edge strength of the pixel is less than the mean edge strength.

25. The storage medium according to claim 24, wherein the generating the binary image of the candidate character strokes includes binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the candidate character strokes by using a window centered at a position of every pixel of the one binary value in the second edge image.

26. The storage medium according to claim 24, wherein the removing the false stroke from the binary image includes:

removing a large stroke by using a width and a height of the stroke.

checking an overlap rate of a contour of a stroke in the binary image of the candidate

character strokes by pixels of the one binary value in the second edge image;

determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold; and

removing the invalid stroke.

27. The storage medium according to claim 24, wherein the binarizing the text line region includes:

determining a size of a window for binarization; and

binarizing a gray scale image of the given image in a Niblack's binarization method to obtain the binary image of the text line region by using the window centered at a position of every pixel of the one binary value in the second edge image.

28. The storage medium according to claim 23, wherein the combining the plurality of strokes into the text line region includes checking whether two adjacent strokes are connectable by using an overlap ratio of heights of the two strokes and a distance between the two strokes, and the combining

the plurality of strokes into the text line region combines the plurality of strokes into a text line region by using a result of checking.

5   29.   The storage medium according to claim 23, wherein the removing the false character stroke from the text line region and reforming the text line region includes:

checking every stroke with a height higher

10   than a mean height of strokes in the text line region;

marking the stroke as a false stroke if the stroke connects two horizontal text line regions into one big text line region;

15   checking every stroke with a width larger than a threshold determined by a mean width of the strokes in the text line region;

marking the stroke as a false stroke if a number of strokes in a region that contains the

20   stroke is less than a predefined threshold; and

reconnecting strokes except for a false stroke in the text line region if the false stroke is detected in the text line region.

25   30.   A computer-readable storage medium storing a

program used to direct a computer, that extracts at least one text line region from a given image, to perform a process comprising:

generating an edge image of the given image;

generating a binary image of candidate character strokes in the given image by using the edge image;

checking an overlap rate of a contour of a stroke in the binary image of the candidate character strokes by pixels indicating an edge in the edge image;

determining that the stroke is a valid stroke if the overlap rate is greater than a predefined threshold and an invalid stroke if the overlap rate is less than the predefined threshold;

removing the invalid stroke; and

outputting information of remaining strokes in the binary image of the candidate character strokes.

31.   A text change frame detection method for selecting a plurality of video frames that includes text contents from given video frames, said method comprising:

removing redundant video frames from the given video frames;

removing video frames that do not contain a text area from the given video frames;

detecting and removing redundant video frames caused by image shifting from the given video frames; and

presenting remaining video frames as candidate text change frames.

32. A text extraction method for extracting at least one text line region from a given image, said method comprising:

generating edge information of the given image;

generating a binary image of candidate character strokes in the given image by using the edge information;

removing a false stroke from the binary image by using the edge information;

combining a plurality of strokes into a text line region;

removing a false character stroke from the text line region and reforming the text line region;

binarizing the text line region by using a height of the text line region; and

presenting a binary image of the text line region.